

工作内容

一、《ScatterNet: A Deep Subjective Similarity Model for Visual Analysis of Scatterplots》

提出背景：

Scagnostics 及其它统计学的特征，能在一定程度上捕捉到散点图的特征，然而，这些特征却与人的感知捕捉(理解)到的散点图特征有一定差距。因此本文提出一种基于神经网络的方法，学习人对散点图的感知特性

方法：

- 1、本文定义了一个三元数据单元，用于训练。分别由一副选定图(anchor 锚图)、相似图、相异图构成。
- 2、通过优化选定图与相似图、相异图之间的欧氏距离来训练一个神经网络。公式如下：

$$Loss_{triplet}(v_{anchor}, v_{pos}, v_{neg}) = \max(0, \|v_{anchor} - v_{pos}\|_2^2 + \alpha - \|v_{anchor} - v_{neg}\|_2^2) \quad (1)$$

- 3、该网络学习到一个关于散点图的特征，可将一幅散点图转化为一个特征向量，该过程实质为 Scatter Embedding 的过程
- 4、将一幅散点图转化为一个特征量，即将两幅散点图的相似性问题，转换为了求两幅散点图的欧式距离的问题

贡献(亮点)：

- 1、本文最大的亮点在于捕捉到了人的感知，即通过设计合理的标注样本方法，进而捕捉到了人对散点图自然地感知，且这份感知具有较大地普遍性(参与测试者仅具有基本地统计知识)。同时也是目前第一篇做深度学习理解可视化地论文
- 2、Evaluation 设计合理，值得为后续工作提供参考

不足：

- 1、使用的神经网络方法较为简单
- 2、研究对象为散点图，属于比较基础的可视化对象

借鉴：

- 1、是否可以研究一种可视化方法，捕捉人类对轨迹相似性的感知，进而做一个 trajectory embedding 的工作？

二、《Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction*》

提出背景：

- 1、现有的方法有 HA（利于均值预测人流量）、ARIMA、SARIMA、VAR、ST-ANN(首个提取时空特征，利用人工神经网络来做预测的)
- 2、本文作者之前的工作为 DeepST
- 3、各相关工作最终预测准确度，其中 ST-ResNet 为本作结果

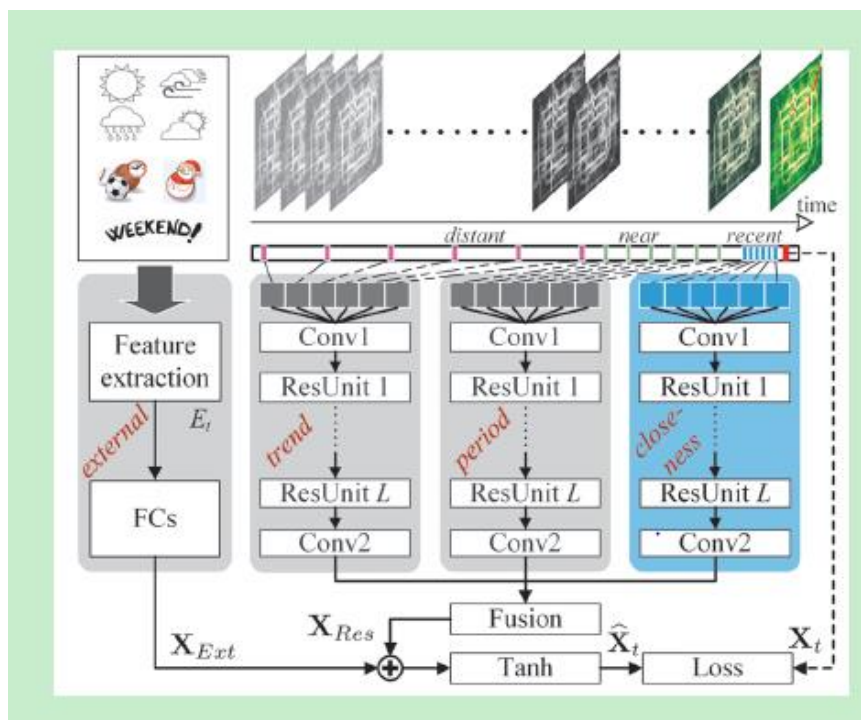
Table 2: Comparison among different methods on TaxiBJ

Model		RMSE
HA		57.69
ARIMA		22.78
SARIMA		26.88
VAR		22.88
ST-ANN		19.57
DeepST		18.18
ST-ResNet [ours]		
L2-E	2 residual units + E	17.67
L4-E	4 residual units + E	17.51
L12-E	12 residual units + E	16.89
L12-E-BN	L12-E with BN	16.69
L12-single-E	12 residual units (1 conv) + E	17.40
L12	12 residual units	17.00
L12-E-noFusion	12 residual units + E without fusion	17.96

4、对于时序数据，还有一种常见的作法 RNNs，而本文使用 CNN

方法：

- 1、从各出租车的轨迹出发，形式化地将轨迹数据转化为一个区域流入和流出流量
- 2、针对问题的特殊性，本文将时间看成一个维度，采用 3 种步长，分段均匀采样地方法，将连续的时间分量离散化，将训练数据转化为了一个三维张量。设计了一个预测网络，借鉴 ResNet(残存网络)结构，使用残存单元改进网络性能，使网络全面地学习到了整幅地图地特征，其网络结构如下：



3、额外综合考虑了天气、工作日/休息日等额外因素的影响，并一并整合到了神经网络中

4、最后输出结果为[-1,1]之间的连续值

贡献：

- 1、提出了一种新的思路处理时序数据的预测问题，通常来说 CNNs 擅于处理图像数据的特征提取，RNNs 处理预测问题
- 2、设计了一种新的网络结构，从结果（提出背景 3）来看，残差单元的使用，对预测的准确度起了很大的作用

借鉴：

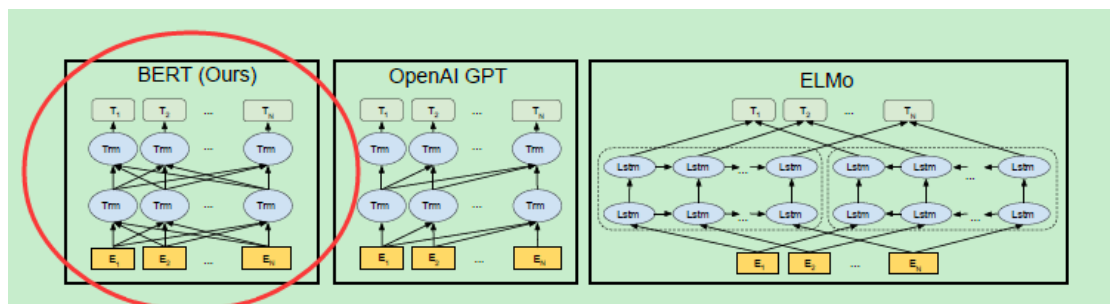
- 1、借鉴这种考虑问题的方法
- 2、使用合理的网络结构(残差单元, 提供更广的感受野), 能极大地提高预测地结果

三、《BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding》

简介: 本文针对 OpenAI GPT 模型中存在的不足, 提出了一个新的模型, 一种新的学习训练方法

贡献:

- 1、本文最大的贡献在于提出了一种新的、半监督的数据标注方法, 从而实现了: 1、更深的上下文联系 2、双向的语言表示 3、前两个特性使得该模型学习到更具泛化性的自然语言特征, 可迁移到其它自然语言处理任务中, 且本文提出的网络结构与问题无关
- 2、模型结构如下图红圈所示



3、训练数据的设计

- Rather than *always* replacing the chosen words with [MASK], the data generator will do the following:
- 80% of the time: Replace the word with the [MASK] token, e.g., my dog is hairy → my dog is [MASK]
- 10% of the time: Replace the word with a random word, e.g., my dog is hairy → my dog is apple
- 10% of the time: Keep the word unchanged, e.g., my dog is hairy → my dog is hairy. The purpose of this is to bias the representation towards the actual observed word.

启示:

- 1、人工智能方向的研究, 训练数据的设计十分关键
- 2、设计的训练数据要与设计的模型结构相契合

四、看完论文后的感想

三篇论文分别从可视化、人流量预测、NLP 迁移学习的角度阐述了各作者对神经网络技术的使用。目前一个大概的想法要做利用神经网络理解交通轨迹可视化结果。这可能与 Visual Question Answering 有所联系，可能需要进一步理解神经网络学习到时空数据的哪些特征，现已找到 CVPR2018 中关于这两个方面的两篇论文，打算接下来进一步阅读

本周工作时间

周一：9 点-8 点 15；scatterNET 看完 1、2 节

周二：9 点至 21 点 20；scatterNET 看到 3.3 节

周三：9 点至 21 点；scatterNET 看到 4.2 节

周四：12 点至 21 点；scatterNET 看到第 5 节之前

周五：9 点至 13 点；看完 scatterNET；看了部分郑宇 AAAI 论文

周六：9 点至 13 点看完郑宇 AAAI 论文，13 点至 21 点看谷歌 BERT 论文

周日：9 点至 21 点，看谷歌 BERT 论文

总结：理论上应该精读完 5 篇论文，时间是足够的。未完成的原因：a)对时间管理得不够 b)刚开始看论文，通读论文花费了过多时间，可以适当地略读部分内容 c)心态还是没有调整好下周应当注意这些方面的问题，力求完成足量的阅读任务